

# 英語スペリング訂正と品詞タグ付けの結合学習

坂口 慶祐<sup>1,a)</sup> 水本 智也<sup>1,b)</sup> 小町 守<sup>1,c)</sup> 松本 裕治<sup>1,d)</sup>

**概要:** 近年, 外国語学習者が書く作文に対する文法の自動誤り訂正が注目を集めているが, 学習者作文の多くは文法的な誤りだけでなくスペリング誤りを多く含んでいる. その結果, 学習者作文に対する品詞タグ付けや構文解析の精度が悪化し, 誤りの訂正を阻害する大きな要因になっている. またスペリング誤り訂正と品詞タグ付けは従来独立したタスクとして扱われており, スペリング誤り訂正の結果が後続の品詞タグ付けや構文解析に影響する点が指摘されてきたが, 近年ではこれまで直列に解析・処理されてきたタスクを統合し, 解析の情報を互いに補完しながら同時に処理する結合学習が盛んになっている. そこで本論文では英語学習者コーパスに対してスペリング誤りと品詞タグ付けの結合学習を行いその効果を検討する. 実験の結果, 結合学習を用いた同時解析の方がそれぞれの解析を単独で行う場合, そしてそれらをパイプラインで処理する場合に比べて解析精度が高くなることを示す.

**キーワード:** 品詞タグ付け, スペリング訂正, 結合学習

## Joint Learning of English Spelling Error Correction and POS Tagging

SAKAGUCHI KEISUKE<sup>1,a)</sup> MIZUMOTO TOMOYA<sup>1,b)</sup> KOMACHI MAMORU<sup>1,c)</sup> MATSUMOTO YUJI<sup>1,d)</sup>

**Abstract:** Automated grammatical error detection and correction tasks for the second language (L2) learners writing of English have become more important in recent years. L2 writing contains not only grammatical errors but also other types of errors such as misspelling and punctuation errors. These ‘ungrammatical’ errors often disturb part-of-speech (POS) tagging and dependency parsing, resulting in an obstacle for grammatical error detection and correction tasks. Spelling error correction and POS tagging have been studied independently but in recent years joint learning of related tasks has been successful in improving NLP pipeline processing. In this paper, we propose a joint learning approach to English spelling error correction and POS tagging. The experimental result shows that the proposed method can correct spelling errors and label POS tags simultaneously for L2 writing as well or better than applying each method independently.

**Keywords:** Part-Of-Speech Tagging, Spelling Error Correction, Joint Learning

### 1. 背景と課題

近年自然言語処理技術の応用先の一つとして, 外国語学習者が書く作文に含まれる文法的な誤りを自動で検出, 訂正するタスクが注目されている. 例えば, 英語学習者作文に対する前置詞の誤り訂正, 冠詞の誤り訂正, 動詞におけ

る時制誤りや一致(人称)誤りの訂正などがある([1], [2], [3], [4]). また英語だけでなく日本語学習者の作文に対する誤り検出, 訂正についても取り組まれている([5], [6]).

一般的にスペリング誤りは文法誤りには含まれず, 文法誤り訂正のタスクにおいてはスペリング誤り訂正を前処理段階で行うか, または特に訂正しないままコーパスを使用することが多い. しかし DeFelice ら [7] が指摘するように, スペリング誤りが文法誤り訂正を阻害する要因の一つであることから, スペリングの誤り訂正は文法誤り訂正に対しても貢献できると考えられるが, 学習者の作文を対象を絞ったスペリング訂正に関する研究は行われていない.

<sup>1</sup> 奈良先端科学技術大学院大学  
Nara Institute of Science and Technology

a) keisuke-sa@is.naist.jp

b) tomoya-m@is.naist.jp

c) komachi@is.naist.jp

d) matsu@is.naist.jp

(例) I think it is \*verey/very \*convent/convenient for the group.  
訂正前: ... ('it', 'PRP'), ('is', 'VBZ'), ('verey', 'PRP'), ('convent', 'NN'), ...  
訂正後: ... ('It', 'PRP'), ('is', 'VBZ'), ('very', 'RB'), ('convenient', 'JJ'), ...

図 1 スペリング誤り訂正情報を用いた品詞タグ付け

Fig. 1 POS tagging using spelling correction.

(例) ... doing social research and some \*analyses/analyses.  
品詞タグ付け結果: ... ('doing', 'VBG'), ('social', 'JJ'), ('research', 'NN'), ('and', 'CC'), ('some', 'DT'), ('analysys', 'NNS'), ...  
訂正候補とその品詞: ['analyses/NNS', 'analysis/NN']

図 2 品詞情報を用いたスペリング訂正

Fig. 2 Spelling correction using POS information.

また従来の自然言語処理研究では、スペリング誤り訂正や品詞タグ付け、構文解析といったタスクは直列的なパイプライン方式で処理されることが一般的であった。つまりスペリング誤り訂正の結果を品詞解析、構文解析の入力とするというように、各レイヤーをボトムアップに処理する方式である。これにはそれぞれのタスクがそれぞれ独立に扱われ、独自に研究されてきたという背景がある。しかし近年の計算機能の向上によって、このようなパイプライン方式で解析・処理されたきたタスクを統合し、同時に処理する結合学習が注目されている。すなわちこれまで独立に扱われていた各タスクの解析に関する情報を、互いに利用しながら統合的に処理するということである。

このような背景から、本論文ではこれまで独立に取り組まれてきたスペリング誤り訂正と品詞タグ付けを統合的に学習し、両タスクを同時に解析する手法を提案する。スペリング誤り訂正と品詞タグ付けを同時に解析することによって、互いの解析情報を利用することができるようになると考えられる。つまり、スペリング誤りを訂正する際には品詞解析の情報を用い、そして品詞タグ付けをする際にスペリング訂正の情報を利用することで、総合的な解析精度を上げることが可能になる。例えば図 1 では、スペリング誤りを訂正することで正しく品詞タグ付けができることを示しており、図 2 は、品詞情報によってスペリング誤り訂正の候補を絞り込むことができることを示している\*1。

以下、第 2 章ではスペリング誤りとその訂正について定義や関連研究について述べる。そして第 3 章では自然言語処理における結合学習について述べ、本稿の提案手法について説明する。第 4 章で今回行った実験について、その手順と結果の説明および考察を行い、まとめと結論を第 5 章で述べる。

## 2. スペリング誤り訂正

スペリング誤りは一般的に書き手の不注意や知識不足により単語の綴りを誤ることであり、

- 誤字 (\*liabrary → library, the → then)
- 同音異義語 (their → there)

などが挙げられる。また Bao ら [8] は、

- 分割 (news paper → newspaper)
- 結合 (soon → so on)

のような例もスペリング誤りとしているが、実施に英語学習者コーパスにおける誤りにもこのような例があることから、本論文でもこれらをスペリング誤りとする。そして本来文法誤りの範疇に含まれる

- 屈折や派生 (showed → shown, like → likely)

に関する誤りも本論文におけるスペリング誤りの対象に含めることとする。なぜなら、本論文が提案する手法ではスペリング訂正と品詞タグ付けを同時に行うため、品詞の違いといった情報もスペリング訂正に活用できると考えられるためである。

次にスペリング誤り訂正に関する関連研究について述べる。スペリング誤り訂正は、どのように誤り箇所を検出し、検出箇所に対して訂正候補を求めて絞り込むかという点が主な問題である。例えば、Chen ら [9] はスペリング誤りを含む検索クエリログに対して、その訂正候補を検索結果から取得する手法を用いており、Islam ら [10] は Google Web 1T コーパスから作成した 3-gram を用いたスペリング誤りの検出・訂正方法を提案している。また Sun ら [11] はウェブ検索のクリックスルーログからスペリング誤りと訂正後のデータ対を取得できることを示した。分割や結合に関するスペリング誤りについては Bao ら [8] によるグラフを用いたアプローチが提案されている。しかしいずれの研究においても、訂正箇所の検出や訂正候補の取得には単語の表層に関する情報のみを扱っており、品詞情報を用いたスペリングの誤り訂正は行われていない。

\*1 これら例文については、代表的な英語学習者コーパスである CLC FCE Dataset から抜粋している。  
<http://ilexir.co.uk/applications/clc-fce-dataset/>  
CLC FCE Dataset の詳細については第 4 章で述べる。また、品詞タグ付けには Natural Language Toolkit (NLTK) を、スペリング誤りの訂正候補出力に GNU Aspell を使用した。  
NLTK 2.0.1 <http://www.nltk.org/>  
GNU Aspell 0.60.6.1 <http://aspell.net/>

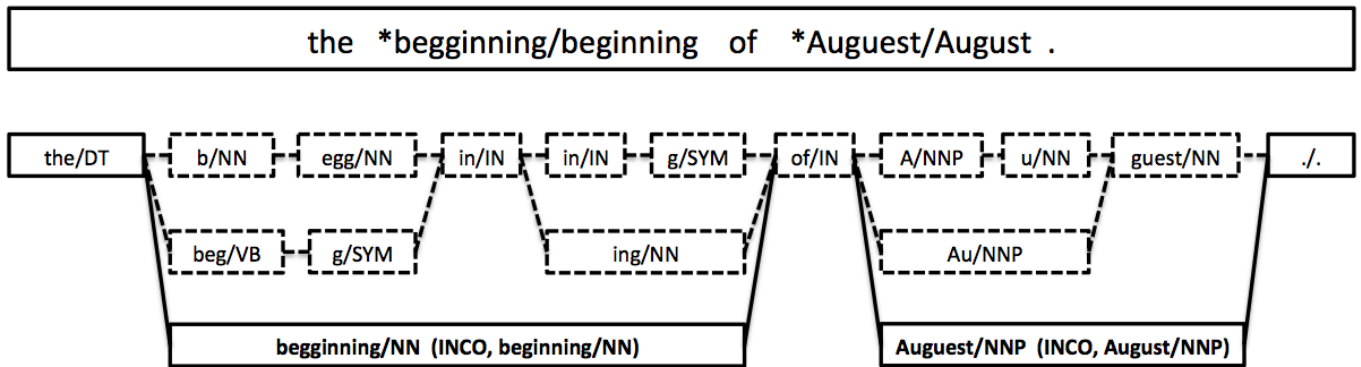


図 3 英語学習者の作文に対する提案手法の適用例  
Fig. 3 Example of the proposed method for L2 writing.

### 3. 自然言語処理における結合学習と提案手法

結合学習とは個々のタスクや処理を1つのモデルに統合し解析する手法であり、近年の計算能力の向上から自然言語処理に適用する研究が増えている。例えば、2008年のCoNLL (Conference on Computational Natural Language Learning) では統語的な依存関係と意味的な依存関係を統合的に学習する Shared Task [15] が開催され、Dahlmeierら [16] は前置詞に対する語義曖昧性解消と意味役割付与を同時に学習する方法を示している。またWatanabeら [17] は述語項構造の大域的な依存関係と述語の語義を結合学習するモデルを提案している。さらに鈴木ら [18] による、日本語の単語、文節、文分割と品詞付与という4つのタスクを同時に解析するという研究も現れている。

いずれの研究においても、結合学習を行った場合の方がそれぞれの解析を直列的に行うパイプライン処理に比べ高い精度を出している。これはそれぞれの解析における情報を互いに補完しながら全体を処理しているためである。ただしこれらの既存研究ではいずれも正しい文に対する同時解析を行っており、本論文が提案するようなスペリング誤りを含んだ文に対する結合学習は研究されていない。

そこで本論文では、各単語に対する品詞のラベル推定とスペリング誤り訂正を同時に解析するモデルを提案する。提案手法は日本語の形態素解析で一般的に用いられるコスト最小法をもとにしている。その際、正しい綴りの単語に加えてスペリング誤りの単語もノードの候補に追加した状態で、品詞の系列ラベリングを行う。ただしスペリング誤りの単語をノードに追加する際には「スペリング誤りがある」と「訂正候補」の情報を持たせておく。これにより品詞タグ付けと同時にスペリングの誤りを訂正することが可能になる。

図 3 を例にとると、\*August/August というスペリング誤りがあった場合、品詞タグ付けの際の単語分割に失敗してしまい、その結果 (A/NNP, u/NN, guest/NN) や

(Au/NNP, guest/NN) のような誤った品詞タグを付与してしまう。しかし事前にスペリング誤りの単語と訂正後の単語を辞書に追加することで、August/NNP (INCO, August/NNP) のように頑健な単語分割および品詞タグ付けを行うと同時に、スペリング誤りである印 (INCO) と訂正候補 (August/NNP) の情報から、誤りの訂正が可能になるという仕組みである。

ここで問題になるのは、訂正候補をどのように取得するかという点である。スペリング誤りの単語の大半は未知語であるため、訂正候補の取得には文字列の類似度 (編集距離) や、スペリング誤り訂正前後の情報を英語学習者コーパスから作成した混同行列を用いる、といった方法が考えられる。これらの方法を用いてスペリング訂正候補を挙げ、単語の生起コストを算出し、それをノードに追加することで、スペリング誤りの単語に対する訂正候補を出しながら品詞タグ付けを行うことが可能になる。また未知語に対する生起コストを決定する際に補関係数入で重み付けをすることで、未知語の生起コストを調整できるようにする。

### 4. 実験

本実験ではスペリング誤りを含んだ文に対して品詞タグの付与とスペリング誤り訂正を同時に行い、その有効性を検証する。

#### 4.1 実験の手順

現在数多くの英語学習者コーパスが作られ、(一部のコーパスについては) 公開されている。しかしながら、いずれの学習者コーパスもデータサイズが小さく、誤りタグに関するポリシーも統一されていないという問題がある。そこで本実験では、英語学習者コーパスの中でも代表的な CLC FCE Dataset \*2 [12] を使用した。

CLC FCE Dataset とは、Cambridge Learners Corpus

\*2 CLC FCE Dataset は The 7th Workshop on Innovative Use of NLP for Building Educational Applications (NAACL-HLT 2012 Workshops) の Shared Task (Preposition and Determiner Error Correction) でも使用されている。

表 1 CLC FCE Dataset における誤りの内訳 (上位 6 件)  
 Table 1 Error Types in the CLC FCE Dataset (Top 6).

誤りタイプ	%	グループ別	%
Spelling	8	Verb	20.8
Replace Verb	6	Punctuation	14.2
Replace Punctuation	6	Spelling	10.7
Tense of Verb	5	Preposition	10.5
Replace Preposition	5	Determiner	9.5
Missing Determiner	5	Noun	9.3

(CLC) から the Cambridge ESOL First Certificate in English (FCE) examination の Writing 試験<sup>\*3</sup>の回答を抽出したコーパスであり、誤りに関するタグが人手で施されている。誤りタグは各品詞やスペリング、句読法などに対してそれぞれ置換 (Replace)、挿入 (Insertion)、削除 (Missing) など誤りの種類に応じたものが施されている<sup>\*4</sup>。

表 1 に CLC FCE コーパスにおける学習者の誤りタイプの上位を示す。この表からもわかるように、学習者の誤りのうち、スペリング誤りが占める割合が大きいことがわかる。そしてこれは CLC FCE コーパスに限ったことではなく、英語学習者の誤りにはある程度同じような傾向があることがわかっている [14]。

実験では CLC FCE コーパスから抽出した 1200 人分のデータファイルを使用した。各データには 1 つまたは 2 つの回答が含まれる。各回答の平均文長は 13.9 文で、平均単語数は 375.8 語である。

品詞の系列ラベリングには CRF (Conditional Random Field) [19] を用い、CRF の学習には汎用的な形態素解析器 MeCab <sup>\*5</sup>を使用した。

学習時に用いるコーパスには訓練データ中の誤りを全て訂正した CLC FCE コーパスに Penn Treebank <sup>\*6</sup>を加えたものを使用した。ただし CLC FCE コーパスには品詞タグが付与されていないため、Penn Treebank および NAIST English Dictionary <sup>\*7</sup> で学習した MeCab <sup>\*8</sup>を用いて、品詞タグを自動付与した。

MeCab で用いる辞書は、

- 表層形
- 左接続状態番号
- 右接続状態番号
- コスト

<sup>\*3</sup> Writing 試験では、エッセイや手紙、校内新聞の記事を書く等の問題が出題される。

<sup>\*4</sup> 誤りタグの詳細については Nicholls [13] を参照のこと。

<sup>\*5</sup> MeCab 0.98  
<http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

<sup>\*6</sup> The Penn Treebank Project Release 2  
<http://www.cis.upenn.edu/~treebank/>

<sup>\*7</sup> NAIST-edict-0.1.0  
<http://sites.google.com/site/masayua/p/naist-edict>

<sup>\*8</sup> 予備実験の結果、MeCab の品詞タグ付与の精度は Precision = 0.944, Recall = 0.944, F-value = 0.944 であった。

```
speaking,37,37,-806,VBG,speak,VB,CORR,*
English:,22,22,0,NNP,English:,NN,CORR,*
absolutely,28,28,-159,RB,absolutely,RB,INCO,absolutely
diffcult,16,16,1480,JJ,diffcult,JJ,INCO,diffcult
```

図 4 MeCab における辞書のフォーマット例  
 Fig. 4 Examples of MeCab dictionary entry.

- 品詞タグ
- 訂正された語の語幹
- 品詞タグ大分類
- 正誤タグ (CORR/INCO)
- 訂正後の単語 (正誤タグが INCO の場合)

を CSV フォーマットで作成したものを使用した。なお図 4 に示される例では MeCab によって左接続状態番号、右接続状態番号、コストを学習した後の値を示している。

学習時に用いる形態素ラティスについては、Penn Treebank コーパスから抽出したものは全て正しいスペリングの単語 (訂正タグが CORR) とし、CLC FCE コーパスでスペリング誤りおよび屈折、派生に関する誤りタグが付いたものに対しては、訂正前の単語 (表層形) と訂正後の単語を取得した混同行列を使用した。

また、1200 ファイルある CLC FCE コーパスからランダムに抽出した 1000 人分のファイルを訓練データとして用い、残りの 200 ファイルのうち 100 ファイルを開発データ、100 ファイルを評価データに使用した。

正解データはスペリング誤りを全て訂正した後、学習コーパス作成時と同様に Penn Treebank および NAIST English Dictionary で学習した MeCab を用いて品詞を自動推定したものをを用いた。<sup>\*9</sup>

また、評価データに含まれる未知語判定、および未知語に対する訂正候補の取得には GNU Aspell を用いて全訂正候補を MeCab のユーザ辞書に追加した。そして訂正候補のコスト算出には Google Web 1T 5-gram Corpus<sup>\*10</sup>から IRSTLM (IRST LM Toolkit)<sup>\*11</sup>を用いて作成した言語モデルを使用した。<sup>\*12</sup> なお未知語の生起コストを決定する際に使用した補間係数  $\lambda$  は開発データから最適値を求めた<sup>\*13</sup>。

## 4.2 評価方法

評価は品詞タグ付けの推定精度およびスペリングの訂正精度を Precision, Recall, F 値を用いて行った。なお品

<sup>\*9</sup> 品詞情報が付与された英語学習者コーパスとして KJ コーパス (Konan-JIEM Learner Corpus) を使用することも考えられるが、本実験で使用した Penn Treebank と品詞体系が異なっているため採用しなかった。

<sup>\*10</sup> Web 1T 5-gram Corpus Version 1.1

<sup>\*11</sup> irstlm 5.70  
<http://sourceforge.net/projects/irstlm/files/irstlm/>

<sup>\*12</sup> ここでは言語モデルで求めた確率  $p$  に対し、 $-\log(p)$  をコストとした。

<sup>\*13</sup> 本実験では  $\lambda=60$  であった。

詞タグ付け, およびスペリング誤り訂正タスクにおける Precision, Recall, F 値は以下のようにして求めた.

- 品詞タグ付けタスク

各品詞ラベルに対して

$$Precision = \frac{\text{正しく品詞付与できた単語数}}{\text{品詞ラベルがついた正解データの単語数}}$$

$$Recall = \frac{\text{正しく品詞付与できた単語数}}{\text{品詞ラベルがついた出力結果の単語数}}$$

$$F\text{-value} = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

- スペリング誤り訂正タスク

$$Precision = \frac{\text{誤りを訂正できた単語数}}{\text{誤り訂正対象となる単語数}}$$

$$Recall = \frac{\text{誤りを訂正できた単語数}}{\text{誤り訂正を行った単語数}}$$

$$F\text{-value} = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

それぞれのタスクにおいて, 品詞タグ付けとスペリング誤り訂正を単独で行った場合とそれらをパイプライン処理した場合, そして提案手法である同時解析を行った場合とで比較を行った. 品詞タグ付けでは学習コーパス作成時に作成した MeCab を使用する条件をベースラインに, またスペリング誤り訂正では GNU Aspell と Google 1T 5gram のみを使用する条件をベースラインとし提案手法との比較を行った.

#### 4.3 実験結果と考察

品詞タグ付けの実験結果は表 2 のようになった. ベースラインと提案手法ではほぼ同等の精度となり, Recall と F 値においてベースラインを上回った. Recall において提案手法が上回った理由はスペリング誤りの単語を分割せず正しく解析できているためと考えられる. 例えば図 5 の例 1 では, (\*anoder/another, DT) という正解に対し, ベースラインでは (an, DT), (oder, NN) と分割して出力したが, 提案手法では (another, DT) のようにスペリング誤り訂正と同時に正しい品詞タグを付与することができていることがわかる. 反対に Precision がベースラインから上がらなかったのは, 訂正候補の品詞が正しく推定できていないためと考えられる. 例えば図 5 の例 2 にあるように (\*arested/arrested, VBN) という正解に対し提案手法では (arrested, VBD) と出力しており, スペリング誤りの訂正は正しいものの, 品詞ラベルを正しく付与することができていないことがわかる. また図 5 の例 3, 例 4 のような分割 (\*some time/sometimes, RB) や結合 (\*Donnot/Do not, VB RB) を含むスペリング誤りに対しては, ベースライン, 提案手法ともに正しく品詞を解析することができていない. そして図 5 の例 5 や例 6 のような派生や屈折を

表 2 品詞タグ付けの結果

Table 2 The result of POS tagging.

	Precision	Recall	F-value
スペリング訂正なし	<b>0.9880</b>	0.9688	0.9788
スペリング訂正あり	0.9879	<b>0.9710</b>	<b>0.9794</b>

表 3 スペリング誤り訂正の結果

Table 3 The result of Spelling error correction.

	Precision	Recall	F-value
品詞タグ付けなし	<b>0.555</b>	0.666	0.605
品詞タグ付けあり	0.489	<b>0.856</b>	<b>0.622</b>

表 4 パイプライン処理との比較

Table 4 Pipeline vs. proposed method.

	Precision	Recall	F-value
パイプライン処理	0.9826	0.9859	0.9842
提案手法	<b>0.9864</b>	<b>0.9887</b>	<b>0.9875</b>

含む誤り (\*please/pleased, JJ), (\*complate/complain, VB) においてもベースライン, 提案手法ともに正しく解析することができていないケースが多かったが, 提案手法では ((\*complate/complain, VB) に対して (complete, VB)) を出力するというようにスペリング訂正は誤ったものの, 正しい品詞タグを付与できた例も見られた. このように分割, 統合, 派生, 屈折誤りにおいて品詞タグ付けのエラーが多かった原因としては, これらの例では訂正前の単語がいずれも存在する正しい単語である点に加え, 本実験で使用した CLC FCE コーパスがエッセーや手紙といったスタイルで書かれている点や, 文内に含まれるその他の文法的な誤りに影響されたことなどが考えられる.

スペリング誤り訂正タスクの結果を表 3 に示す. 実験の結果, 提案手法はベースラインに比べ Recall で大きく上回ったが, Precision は下がる結果となった. スペリング誤り訂正に成功した例としては, 図 6 の (\*descide/decide) のなどがあり, ベースラインで正しく訂正できているものは提案手法においても同様に訂正できているものが多かった. ベースラインに比べ Recall が上回り, Precision が下がった原因としては, 提案手法では訂正前の単語が正しい単語として存在している場合に対してもスペリング誤り訂正検出が可能である一方, 結果的に正しい形には訂正できなかったためと考えられる. 図 6 の (\*than/then) を例にとると, スペリング訂正対象となる than は実際に存在する単語であり, ベースラインでは訂正が行われない. しかし提案手法では, 品詞タグの情報をを用いることで表層的に正しい単語に対してもスペリング誤りを検出することができるため, このような訂正が行われたと考えられる.

最後にパイプライン処理を行った場合と提案手法との品詞タグ付け精度の比較について述べる. 表 4 にあるよう

図 5 品詞タグ付け結果の例

Fig. 5 Examples of POS tagging result.

例 1:	Will b good for yor group you don't need move to <u>anoder</u> place.
正解データ:	… ('move', 'VB'), ('to', 'TO'), ('another', 'DT'), ('place', 'NN'), …
ベースライン:	… ('move', 'VB'), ('to', 'TO'), ( <u>'an', 'DT'</u> ), ('oder', 'NN'), ('place', 'NN'), …
提案手法:	… ('move', 'VB'), ('to', 'TO'), ( <u>'another', 'DT'</u> ), ('place', 'NN'), …
例 2:	Then police realized who were bank robbers and <u>arested</u> them.
正解データ:	… ('bank', 'NN'), ('robbers', 'NNS'), ('and', 'CC'), ( <u>'arrested', 'VBD'</u> ), ('them', 'PRP'), …
ベースライン:	… ('bank', 'NN'), ('robbers', 'NNS'), ('and', 'CC'), ( <u>'a', 'DT'</u> ), ( <u>'rested', 'VBD'</u> ), ('them', 'PRP'), …
提案手法:	… ('bank', 'NN'), ('robbers', 'NNS'), ('and', 'CC'), ( <u>'arrested', 'VBN'</u> ), ('them', 'PRP'), …
例 3:	<u>Now</u> a day clothes is designed to be more comfortable.
正解データ:	( <u>'Nowadays', 'NNS'</u> ), ('clothes', 'NNS'), ('is', 'VBZ'), …
ベースライン:	( <u>'Now', 'RB'</u> ), ( <u>'a', 'DT'</u> ), ('day', 'NN'), ('clothes', 'NNS'), ('is', 'VBZ'), …
提案手法:	( <u>'Now', 'RB'</u> ), ( <u>'a', 'DT'</u> ), ('day', 'NN'), ('clothes', 'NNS'), ('is', 'VBZ'), …
例 4:	<u>Donnot</u> forget to wear a pair of comfortable shoes, because …
正解データ:	( <u>'Do', 'VBP'</u> ), ( <u>'not', 'RB'</u> ), ('forget', 'VB'), ('to', 'TO'), …
ベースライン:	( <u>'Don', 'NNP'</u> ), ( <u>'not', 'RB'</u> ), ('forget', 'VB'), ('to', 'TO'), …
提案手法:	( <u>'Don', 'NNP'</u> ), ( <u>'not', 'RB'</u> ), ('forget', 'VB'), ('to', 'TO'), …
例 5:	I was very <u>please</u> when I knew …
正解データ:	… ('was', 'VBD'), ('very', 'RB'), ( <u>'pleased', 'JJ'</u> ), ('when', 'WRB'), ('I', 'PRP'), …
ベースライン:	… ('was', 'VBD'), ('very', 'RB'), ( <u>'pleased', 'VB'</u> ), ('when', 'WRB'), ('I', 'PRP'), …
提案手法:	… ('was', 'VBD'), ('very', 'RB'), ( <u>'pleased', 'VB'</u> ), ('when', 'WRB'), ('I', 'PRP'), …
例 6:	I am writing to you to <u>complate</u> about the show …
正解データ:	… ('you', 'PRP'), ('to', 'TO'), ( <u>'complain', 'VB'</u> ), ('about', 'IN'), ('the', 'DT'), …
ベースライン:	… ('you', 'PRP'), ('to', 'TO'), ( <u>'com', 'NN'</u> ), ( <u>'plate', 'NN'</u> ), ('about', 'IN'), ('the', 'DT'), …
提案手法:	… ('you', 'PRP'), ('to', 'TO'), ( <u>'complete', 'VB'</u> ), ('about', 'IN'), ('the', 'DT'), …

図 6 スペリング誤り訂正のエラー分析

Fig. 6 Error analysis of of Spelling correction.

例:	… we all agree to listen to music so <u>than</u> we <u>descide</u> to come to your musical show …
正解データ:	… we all agree to listen to music so <u>then</u> we <u>decide</u> to come to your musical show …
ベースライン:	… we all agree to listen to music so <u>than</u> we <u>decide</u> to come to your musical show …
提案手法:	… we all agree to listen to music so <u>that</u> we <u>decide</u> to come to your musical show …

に、提案手法<sup>\*14</sup>が Precision, Recall, F 値全てにおいてパイプライン処理を上回る結果となった。これはパイプライン処理では独立して扱われたスペリング訂正と品詞の情報が、提案手法では統合的に用いられたためと考えられる。

## 5. 結論

本論文では、これまで独立に解析されてきたスペリング誤り訂正と品詞タグ付けの結合学習を提案した。実験の結果、スペリング誤り訂正、品詞タグ付けのいずれにおいても、個別に処理するよりも Recall, F 値が上回る結果となった。またパイプライン処理との比較においては、Precision,

Recall, F 値の全てで提案手法がパイプライン処理を上回る解析精度であることが示された。これはスペリング誤り訂正と品詞タグ付けの解析結果を統合的に用いることで、それぞれの精度が上がるためだと考えられる。また分割・結合、派生・屈折に関する誤りのように、訂正前の単語が実際に存在する単語である場合、誤りの検出はできるものの正しく訂正できない事例が多いことを確認した。このような場合におけるスペリング誤り訂正と品詞タグ付けについては今後の課題としたい。

<sup>\*14</sup> パイプライン処理での未知語の訂正候補に合わせ、提案手法においても全訂正候補を辞書に追加するのではなく、1-best のみを訂正候補として辞書に追加した MeCab を用いた。このため表 2 の結果とは値が異なっている。

## 参考文献

- [1] Rozovskaya A. and Roth D.: Algorithm selection and model adaptation for ESL correction tasks, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, 924-933 (2011).
- [2] Nagata R., Wakana T., Masui F., Kawai A. and Isu N.: Detecting article errors based on the mass count distinction, *Proceedings of the Second international joint conference on Natural Language Processing*, 815-826 (2005).
- [3] John L. and Stephanie S.: Correcting Misuse of Verb Forms, *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics*, 174-182 (2008).
- [4] 田尻 俊宗, 小町 守, 松本 裕治: 大域的文脈情報を用いた英語時制誤りの検出と訂正, 言語処理学会第 18 回年次大会発表論文集 357-360 (2012).
- [5] Mizumoto T., Komachi M., Nagata M., and Matsumoto Y.: Mining Revision Log of Language Learning SNS for Automated Japanese Error Correction of Second Language Learners, *Proceedings of 5th International Joint Conference on Natural Language Processing*, 147-155 (2011).
- [6] 笠原 誠司, 藤野 拓也, 小町 守, 永田 昌明, 松本 裕治: 日本語学習者の誤り傾向を反映した格助詞訂正, 言語処理学会第 18 回年次大会発表論文集 14-17 (2012).
- [7] De Felice R. and Pulman S.: Automatic Detection of Preposition Errors in Learner Writing, *Calico In Calico (The Computer Assisted Language Instruction Consortium)*. Vol. 26. No. 3. 512-528. (2009).
- [8] Bao Z., Kimelfeld B, and Li Y.: A Graph Approach to Spelling Correction in Domain-Centric Search, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. 905-914 (2011).
- [9] Chen Q., Li M., and Zhou M.: Improving Query Spelling Correction Using Web Search Results, *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*. 181-189. (2007).
- [10] Islam A. and Inkpen D.: Real-word spelling correction using Google Web IT 3-grams, *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* . 1241-1249 (2009).
- [11] Sun X., Gao J., Micol D., and Quirk C.: Learning phrase-based spelling error models from clickthrough data, *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. 266-274 (2010).
- [12] Yannakoudakis H., Briscoe T., and Medlock B.: A New Dataset and Method for Automatically Grading ESOL Texts, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. 180-189 (2011).
- [13] Nicholls D.: The Cambridge Learner Corpus - error coding and analysis for lexicography and ELT, *Summer Workshop on Learner Corpora*. 572-581 (2003).
- [14] Leacock C., Chodorow M., Gamon M., and Tetreault J., *Automated Grammatical Error Detection for Language Learners*. Morgan and Claypool Publishers. (2010).
- [15] Surdeanu M., Johansson R., Meyers A., Márquez L., and Nivre J., Shared Task on Joint Parsing of Syntactic and Semantic Dependencies, *Proceedings of the Twelfth Conference on Natural Language Learning*. 159-177 (2008).
- [16] Dahlmeier D., Ng H.T. and Schultz T.: Joint learning of preposition senses and semantic roles of prepositional phrases, *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 450-458. (2009).
- [17] Watanabe Y., Asahara M. and Matsumoto Y.: A structured model for joint learning of argument roles and predicate senses, *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, 98-102 (2010).
- [18] 鈴木潤, Duh Kevin, 永田昌明: 拡張ラグランジュ緩和を用いた同時自然言語解析法, 言語処理学会第 18 回年次大会発表論文集 1284-1287 (2012).
- [19] Lafferty J., McCallum A. and Pereira F.: Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data, *Proceedings of the Eighteenth International Conference on Machine Learning*, 282-289 (2001).